Application of hierarchical Bayesian models to extreme precipitation events

Kári Hlynsson

1 Introduction

- Describe the pretext of the project; application to protection of infrastructure, prevention of disasters etc.
- Why is the Bayesian framework useful for the modelling of precipitation phenomena?
- What are the prior attempts of modelling, how are they modelled. Is this project a novel approach in any way?

2 Preliminaries

2.1 Theoretical background

- The generalised extreme value distribution
  - Parameters
  - Support, cover for different values of ξ
  - Describe usual ranges of parameters in the context of environmental modelling, especially precipitation – cf. literature
- Fisher-Tippett theorem states that block maxima of an i.i.d. collection of random variables converges to the GEV distribution
- Return levels?

2.2 Markov-Chain Monte Carlo (MCMC) sampling and `stan`

- Describe the general idea underlying MCMC sampling, i.e. the high-dimensional integral setting involved in estimation
- Hamiltonian Monte Carlo and the NUTS algorithm.

2.3 Model diagnostics and evaluation

- Talk about $\widehat{\text{ESS}}$, $\hat{R}$, autocorrelation and trace plots.
- Mention how I will evaluate convergence of chains.
- Describe LOO, PSIS, and WAIC

3 Bayesian modelling framework

3.1 Fixed parameter models

- Null model $Y_n \sim \text{GEV}(\mu, \sigma, \xi)$ where $\mu \sim \mathcal{N}(0, \sigma_\mu^2)$, $\sigma \sim \text{Exp}(\lambda_\sigma)$, and $\xi \sim \text{Beta}(\alpha_\xi, \beta_\xi)$ transformed onto the interval $[-0.5, 0.5]$.

3.2 Timetrend model

- Timetrend model $Y_n \sim \text{GEV}(\mu_n, \sigma, \xi)$ where $\mu_n = \mu_0\big(1 + \Delta(t - t_c)\big)$, $\Delta \sim \mathcal{N}(0, \sigma_\Delta^2)$, $\sigma \sim \text{Exp}(\lambda_\sigma)$ and $\xi \sim \text{Beta}(\alpha_\xi, \beta_\xi)$ shifted onto the interval $[-0.5, 0.5]$. Ω

3.3 Hierarchical model

- Split the time interval $\{1, \ldots, T\}$ into m evenly spaced points excluding time T, $1 = t_1 < t_2 < \ldots < t_m < T$.
- We let $\mu_t = \sum_{j=1}^{k} \beta_j (t - t_j)\mathbf{1}_{\{t \geq t_j\}}$, assign priors $\sigma \sim \text{Exp}(\lambda_\sigma)$ and $\xi \sim \text{Beta}(\alpha_\xi, \beta_\xi)$ shiften onto the interval $[-0.5, 0.5]$.
- Three priors for the vector $\boldsymbol{\beta}$ are compared, namely
  - an i.i.d. Gaussian noise prior, $\beta_j \sim \mathcal{N}(0, \sigma_\varepsilon^2)$,
  - a random walk prior, $\beta_1 \sim \mathcal{N}(0, \sigma_\beta)$ and $\beta_j \sim \mathcal{N}(\beta_{j-1})$
  - an AR1 prior where $\beta_1 \sim \mathcal{N}(0, \sigma_\beta)$ and $\beta_j \sim \mathcal{N}(\alpha + \phi\beta_{j-1}, \sigma_\varepsilon)$ for $j \geq 2$
- We then optimize over the hyperparameter m to obtain the best results

4 Results

- Mention convergence good excluding some divergent transitions (proportion less than 3%), trace plots are shown in appendix
- $\hat{R} \approx 1$, $\widehat{\text{ESS}} > N$
- Table with parameter estimate tables for each of the parameters
- Table with WAIC, LOO, and PSIS for each of the models side by side with $m = 6$ as candidate
- Table showing WAIC, LOO, and PSIS and a figure showing trends of these estimates as a function of $m \in [n]$

Tab. 1. Parameter summary statistics for the null model.

	Mean	Mean SE	SD	2.5%	50%	97.5%	$\widehat{\text{ESS}}$
μ	25.3575	0.0046	0.7484	23.9136	25.3489	26.8523	26294
σ	7.7143	0.0033	0.5395	6.7254	7.6897	8.8425	26309
ξ	-0.0106	0.0004	0.0604	-0.1222	-0.0130	0.1134	23787

Tab. 2. Parameter summary statistics for the timetrend model.

	Mean	Mean SE	SD	2.5%	50%	97.5%	$\widehat{\text{ESS}}$
μ_0	25.3093	0.0053	0.7164	23.9188	25.2998	26.7283	18283
σ	7.2481	0.0041	0.5370	6.2669	7.2217	8.3736	17497
ξ	0.0267	0.0005	0.0670	-0.0979	0.0247	0.1630	17907
Δ	0.0022	0.0000	0.0006	0.0009	0.0021	0.0034	49770

Tab. 3. Parameter summary statistics for the i.i.d. hierarchical model.

	Mean	Mean SE	SD	2.5%	50%	97.5%	$\widehat{\text{ESS}}$
β_1	2.2298	0.0018	0.2277	1.7759	2.2336	2.6718	15951
β_2	−3.2579	0.0049	0.5961	−4.4262	−3.2634	−2.0913	14792
β_3	1.7557	0.0052	0.6593	0.5108	1.7382	3.0974	16318
β_4	−1.1203	0.0045	0.6225	−2.3156	−1.1303	0.1306	19194
β_5	0.6396	0.0045	0.6276	−0.5978	0.6374	1.8775	19537
β_6	−0.2340	0.0043	0.5988	−1.3858	−0.2476	0.9747	19817
β_7	0.2380	0.0042	0.6065	−0.9291	0.2269	1.4593	20673
β_8	−0.8775	0.0042	0.6611	−2.1926	−0.8649	0.3987	24304
σ	9.0816	0.0058	0.8057	7.6243	9.0454	10.7854	19423
τ	2.0024	0.0096	2.0019	0.0482	1.4161	7.3918	43555
ξ	0.0019	0.0006	0.0828	−0.1502	−0.0017	0.1730	16960

Tab. 4. Parameter summary statistics for the random walk hierarchical model.

	Mean	Mean SE	SD	2.5%	50%	97.5%	$\widehat{\text{ESS}}$
β_1	1.8779	0.0026	0.2753	1.3014	1.8884	2.3879	11368
β_2	-2.2532	0.0066	0.6797	-3.5642	-2.2665	-0.8758	10743
β_3	0.7112	0.0059	0.6394	-0.4595	0.6915	2.0363	11680
β_4	-0.5104	0.0042	0.5293	-1.5728	-0.4988	0.5064	16151
β_5	0.3131	0.0038	0.4908	-0.6317	0.3068	1.2964	16702
β_6	-0.0404	0.0035	0.4682	-0.9591	-0.0443	0.8986	17806
β_7	0.0611	0.0036	0.4786	-0.8490	0.0471	1.0383	17755
β_8	-0.6947	0.0039	0.5661	-1.8410	-0.6806	0.3889	21513
σ	8.8751	0.0057	0.8147	7.3887	8.8358	10.5877	20115
τ	2.3623	0.0078	0.9292	0.9751	2.2162	4.5857	14083
ξ	0.0421	0.0007	0.0896	-0.1244	0.0388	0.2261	16188

Tab. 5. Parameter summary statistics for the AR1 hierarchical model.

	Mean	Mean SE	SD	2.5%	50%	97.5%	$\widehat{\text{ESS}}$
beta[1]	1.9853	0.0021	0.2310	1.5324	1.9853	2.4421	12600
beta[2]	-2.6416	0.0052	0.5657	-3.7952	-2.6304	-1.5543	11657
beta[3]	1.2968	0.0051	0.5710	0.2223	1.2777	2.4655	12321
beta[4]	-1.0662	0.0044	0.5231	-2.0865	-1.0760	-0.0197	14002
beta[5]	0.7417	0.0044	0.5170	-0.2772	0.7404	1.7540	14093
beta[6]	-0.3992	0.0042	0.5014	-1.3911	-0.4043	0.5987	14351
beta[7]	0.3087	0.0042	0.5227	-0.6862	0.2950	1.3778	15596
beta[8]	-0.7093	0.0040	0.5497	-1.8376	-0.6958	0.3341	19044
sigma	8.7397	0.0064	0.7942	7.3026	8.6966	10.4252	15177
tau	0.8126	0.0030	0.3545	0.3718	0.7363	1.6891	14176
phi	-0.7695	0.0020	0.2454	-1.2245	-0.7825	-0.2521	15359
xi	0.0491	0.0008	0.0897	-0.1168	0.0456	0.2329	11860

5 Discussion

- Figure of predicted values with 95% credible sets against observed values for best model

A Convergence of MCMC samplers

B Code used in the project